JOURNAL OF COMBINATORIAL THEORY 2, 507-513 (1967)

On Comparing Connecting Networks

VACLAV E. BENEŠ

Bell Telephone Laboratories, Incorporated,
Murray Hill, New Jersey

ABSTRACT

Seeing how *structural* differences between connecting networks lead to differences in their *performance* is a basic problem in telephone traffic theory. The object is to transform combinatorial information about networks into an inequality between suitable blocking probabilities. This paper stresses the relevance of routing to this problem, and takes an initial step toward answering the equestion: What kinds of relationships between two networks ensure that one is "better" than the other?

A relation $\leq$ is defined which partially orders all the possible networks ν on given inlets and outlets. With an *assignment* defined as a specification of what inlet is to be connected to what outlet, $\nu_1 \leq \nu_2$ means roughly that it is possible to map a subset of the states of ν_1 that is closed under hangups onto those of ν_2 so as to preserve assignments, and in such a way that only states comparable in the natural partial ordering can have comparable images.

With $b(\nu, R)$ the probability of blocking of network ν under routing rule R (appropriate to ν), it is proved (i) that $\min_R b(\nu, R)$ is isotone on $\leq$, and (ii) that $\nu_1 \leq \nu_2$ implies the existence of an isomorph of the states of ν_2 within ν_1. The latter result, suggested by S. Darlington, provides a different, very natural proof of the isotony (i). The intuitive meaning of these two results is that, if $\nu_1 \leq \nu_2$, then any way of operating ν_2 can be mimicked in ν_1, so that the best way of routing in ν_1 gives a loss no greater than that achieved by the best way of routing in ν_2.

1. INTRODUCTION AND SUMMARY

In the design of connecting networks it is customary to compare alternative networks by estimating their respective carried loads and loss probabilities when subjected to the same traffic sources.[1] In practice such comparisons are often undertaken on a pairwise basis and carried out numerically. Two more or less promising networks are selected and analyzed for load and loss by traffic simulation on a digital computer. The networks chosen usually differ significantly in point of structure, switch size, or routing, and the aim of the calculation is to plot the effects of these differences on performance, as measured by the probability of loss. We shall not be concerned with this kind of comparison here. Rather, taking the view that in traffic theory it is desirable to have some general or "wholesale" methods, we try to provide some of the more elementary answers to the question: What kinds of relationships between two networks ensure that one is "better" than the other in point of loss?

To this end we describe a way of comparing networks which proceeds by transforming *combinatorial information* about networks into an *inequality* between suitable blocking probabilities. Specifically, we let $b(\nu, R, \lambda)$ be the probability of blocking for network ν operated according to routing rule R at traffic level λ; we then partially order all the possible networks on given inlets and outlets by a special relation $\leq$, and we show that for all $\lambda > 0$

$$\beta(\nu, \lambda) = \min_R b(\nu, R, \lambda)$$

is *isotone* on the partial ordering $\leq$, *i.e.*,

$$\nu_1 \leq \nu_2 \quad \text{implies} \quad \beta(\nu_1, \lambda) \leq \beta(\nu_2, \lambda).$$

The heuristic interpretation of "$\nu_1 \leq \nu_2$" is that (in a precise sense to be given) ν_1 is "at least as good" as ν_2. This interpretation is justified by two results: (1) by the isotony of $\beta(\cdot, \lambda)$, i.e., by the fact that the best way of operating ν_1 gives a probability of blocking no larger than that achieved by the best way of operating ν_2; and (2) by the fact that, with $S(\nu)$ the set of states of ν, $\nu_1 \leq \nu_2$ implies that there is an isomorph of $S(\nu_2)$ within $S(\nu_1)$ on which any sequence of events in ν_2 can be mimicked. The possibility of (2) was suggested by S. Darlington and leads to a direct proof of the isotony of $\beta(\cdot, \lambda)$ not based on the routing theory in reference 1.

---

[1] Two parameters, the carried load and the loss, must be used in precise studies because the finite source effect will affect each network in a different way, so that attempt rates in the two will differ; as larger systems are considered and the traffic per line becomes negligible, the difference vanishes. This is the passage to infinite sources, in which carried loads become accurate guides to the loss.

Our approach is based on the following remarks:

1. It is doubtful whether all the networks that can be placed between a given set of terminals can be ordered in an array remotely resembling a linear one; it is rather more likely that there are several useful *partial* orderings of these networks. Such an ordering will be described.

2. For some pairs of networks it is probably true that, if one network is "better" than the other (say in loss incurred by the same sources) for one traffic value, then it is better at all traffic values. A full understanding of the conditions under which this phenomenon occurs would facilitate the comparison of networks by removing the dependence of the comparison on traffic level, and making the comparison an essentially combinatorial matter. Our results shed light on this problem because the comparisons we obtain are independent of the traffic offered.

3. Comparing networks is greately complicated by the effect of routing rules on system performance. What is actually compared in practice is one network ν_1, and a rule R_1 for operating ν_1, against another network ν_2 with a rule R_2 for running ν_2. However, ideally one should try to compare the *best* way of operating ν_1 with the *best* way of operating ν_2, for only then do the networks show their full combinatorial potential. Thus the comparison of networks with respect to performance cannot and should not be divorced from considerations of optimal routing. The methods and concepts used in a previous study [1] of routing will be applied to this aspect of network comparison.

4. Finally, let us consider a case in which a network ν_1 differs from another ν_2 only in that some switching equipment is present in ν_1 and absent in ν_2; for example, ν_1 might be a three-stage Clos network [2], and ν_2 might be obtained from it by removing one of the middle switches. In this and many similar cases we would be predisposed to declare at once that ν_1 is "better" than ν_2, and that it is "intuitively obvious" that, if each network is operated according to an optimal policy, then ν_1 yields a blocking probability no greater than that of ν_2. Yet how is one to prove this and similar "obvious" facts? The methods to be described here provide an answer.

2. Preliminaries

Notations and terminology from Reference 1 will be used. In addition we define the gain function $g(\cdot,\cdot)$ for $x \in S$ and $c \in x$ by

$$g(c, x) = \begin{cases} 1, & \text{if } c \text{ is not blocked in } x, \\ 0, & \text{if } c \text{ is blocked in } x. \end{cases}$$

We now define the set $N(I, \Omega)$ of all possible networks $\nu = \{G, I, \Omega, S\}$ for which the set I of inlets and the set Ω of outlets are fixed, but the graph G depicting network structure and the set S of states may vary in any way consistent with their defining a network in the sense of reference 2.

A set $X \subseteq S$ is *closed from below* if $x \in X$, $y \leq x$ imply $y \in X$.

For $\nu \in N(I, \Omega)$, let $C(\nu)$ be the set of all possible routing matrices [1] for the network ν.

$N(I, \Omega)$ is partially ordered by the following relation $\leq$:

$\nu_1 \leq \nu_2 \equiv \exists$ domain $D \subset S(\nu_1)$, D closed from below, and $\exists$ onto map μ: $D \to S(\nu_2)$ such that

(i) μ preserves assignments: $\gamma(\mu x) = \gamma(x)$, and

(ii) $x, y \in D$, $\mu x \geq \mu y$ imply $x \geq y$.

In the following three lemmas, μ is to be a map satisfying the conditions defining $\leq$.

LEMMA 1. *μ preserves dimension*: $|\mu x| = |x|$.

PROOF: $|\mu x| = |\gamma(\mu x)| = |\gamma(x)| = |x|$.

LEMMA 2. *$x \in D$, $c \in x$ imply* $A_{c(\mu x)} \subseteq \mu(A_{cx} \cap D)$.

PROOF: Take $y \in A_{c(\mu x)}$. Since μ is onto, $y = \mu w$ for some $w \in D$. But $\gamma(w) = \gamma(y) = \gamma(\mu x) \cup c \cup \gamma(x) \cup c$. Hence, by (ii), $w \in A_{cx} \cap D$.

LEMMA 3. *μ diminishes gain, i.e., $x \in D$, $c \in x$ imply*

$$g(c, x) \geq g(c, \mu x),$$

so every call blocked in x is blocked in μx.

PROOF: If c is blocked in x, $A_{cx} = \varphi$; hence Lemma 2 gives

$$A_{c(\mu x)} \subseteq \mu(A_{cx} \cap D) \subseteq \varphi.$$

3. ISOTONY THEOREM

We shall compare networks by using the natural criterion most widely accepted by telephone engineers as a suitable measure of performance for a connecting network, the probability of blocking, Pr{bl}. Suitable definitions of this quantity appear in References 1 and 2. According to Theorem 4 of Reference 1, this probability is smallest if the fraction of events that are successful calls is largest. Let the traffic parameter λ be fixed henceforth, and set $b(\nu, R) = \Pr(\mathrm{bl}\}$ for ν operated according to policy R

THEOREM 1. *If* $\nu_1 \leq \nu_2$, *then*

$$\min_{R \in C(\nu_1)} b(\nu_1, R) \leq \min_{R \in C(\nu_2)} b(\nu_2, R)$$

PROOF: Since $\nu_1 \leq \nu_2$, there exists a map μ, with domain $D \subseteq S(\nu_1)$, satisfying (i)–(iii). By Theorems 4 and 15 of reference 1, it is enough to show that $x \in D$ implies

$$E_x(n) \geq E_{\mu x}(n), \qquad n = 1, 2, ...,$$

where

$E_x(n) =$ expected number of successful call attempts in n events, starting in state x, and following an optimal policy.

Lemma 1 gives $| \mu x | = | x |$, and since $\gamma(x) = \gamma(\mu x)$, it can be seen that $\alpha_x = \alpha_{\mu x}$. Lemma 3 implies $s(x) \geq s(\mu x)$, and so

$$E_x(1) = \frac{\lambda s(x)}{| x | + \lambda \alpha_x} \geq \frac{\lambda s(\mu x)}{| \mu x | + \lambda \alpha_{\mu x}} = E_{\mu\lambda}(1).$$

As a hypothesis of induction assume that $x \in D$ implies $E_x(n) \geq E_{\mu x}(n)$. For $c \in x$ Lemma 2 gives

$$\begin{aligned} \max_{y \in A_{cx}} E_y(n) &\geq \max_{y \in A_{cx} \cap D} E_y(n) \\ &\geq \max_{y \in A_{cx} \cap D} E_{\mu y}(n) \\ &\geq \max_{z \in \mu(A_{cx} \cup D)} E_z(n) \\ &\geq \max_{z \in A_{c(\mu x)}} E_z(n). \end{aligned} \tag{1}$$

Now

$$\begin{aligned}
E_x(n+1) &= \frac{\lambda}{|x|+\lambda\alpha_x} \sum_{\substack{c\in x \\ c \text{ not blocked in } x}} \max\{g(c,x)+ \max_{y\in A_{cx}} E_y(n), E_x(n)\} \\
&\quad + \frac{\lambda}{|x|+\lambda\alpha_x} \sum_{\substack{c\in x \\ c \text{ blocked in } x}} E_x(n) + \frac{1}{|x|+\lambda\alpha_x} \sum_{h\in x} E_{x-h}(n) \\
&\geq \frac{\lambda}{|\mu x|+\lambda\alpha_{\mu x}} \sum_{\substack{c\in x \\ c \text{ not blocked in } x}} \max\{g(c,x)+ \max_{y\in A_{cx}} E_y(n), E_x(n)\} \\
&\quad + \frac{\lambda}{|\mu x|+\lambda\alpha_{\mu}} \sum_{\substack{c\in x \\ c \text{ blocked in } \mu x}} E_\lambda(n) + \frac{1}{|\mu x|+\lambda\alpha_{\mu x}} \sum_{h\in \mu x} E_{x-h}(n) \\
&\geq E_{\mu x}(n+1),
\end{aligned}$$

by (1) and the induction hypothesis, using the fact that D is closed under hangups. This proves the theorem.

4. Darlington's Direct Argument

In conversation, S. Darlington has asked whether $\nu_1 \leq \nu_2$ implies that it is possible to embed $S(\nu_2)$ within $S(\nu_1)$, and thus to mimic in ν_1 any sequence of events in ν_2. He pointed out that if this is so then any routing rule R_2 for ν_2 would have a direct analog R_1 in ν_1 which would give at least as good a performance for ν_1 as R_2 provided for ν_2; the isotony theorem could then be proved in a natural intuitive way without recourse to the theory of reference 1. We now show that Darlington's program can be carried out.

A precise meaning of "mimicry of ν_2 within ν_1" is provided by the concept of *isomorphism* [3]. An isomorphism between two partially ordered systems is a one-to-one correspondence that preserves order in both directions. An isomorph of $S(\nu_2)$ in $S(\nu_1)$ would be a subset M of $S(\nu_1)$ and a correspondence i: $M \leftrightarrow S(\nu_2)$ such that $x \geq y$ if and only if $ix \geq iy$.

THEOREM 2. *If $\nu_1 \leq \nu_2$, then there exist a domain $M \subseteq S(\nu_1)$ and a correspondence i: $M \leftrightarrow S(\nu_2)$ such that $x, y \in M$ implies $\gamma(ix) = \gamma(x)$, and $x \geq y$ if and only if $ix \geq iy$.*

PROOF: Since $\nu_1 \leq \nu_2$ there exist a domain $D \subseteq S(\nu_1)$ and a map μ satisfying conditions (i) and (ii) in the definition of $\leq$ between networks. For each maximal y of ν_2 choose one state from $\mu^{-1}(\{y\})$, and let M be the closure from below (i.e., under hangups) of the states of ν_1 so chosen. Define i to be the restriction of μ to M. It is obvious that i preserves assignments; it suffices to show that (1) i is one-to-one, (2) $i(M) = \mu(D)$, and (3) $x \geq y$ iff $ix \geq iy$, on M.

$i \subseteq \mu$, so i is a function. If $ix = iy$ for $x \neq y$, $x, y \in M$, then ix is not maximal. Let μz, $z \in D$, cover ix, hence iy, so that z covers x and y, and $x, y \in B_z$. But $\gamma(x) = \gamma(ix) = \gamma(y)$. Since B_z cannot contain more than one state realizing a given assignment, $x = y$, and (1) is proved.

If $x \in S(\nu_2)$ is maximal, then $x \in i(M)$. Let $y \,\varepsilon\, S(\nu_2)$ be a state of largest possible dimension that does not belong to $i(M)$. Then, for some $x \in M$, $y \in B_{ix}$, and so, for some $z \in B_x$, $y = \mu z$. Since M is closed from below, $z \in M$, $iz \in i(M)$, and $y \in i(M)$. This proves (2).

Since $i \subseteq \mu$, it is clear that $x, y \in M$, $ix \geq iy$ imply $x \geq y$. Take then $x, y \in M$ with $y \in B_x$. Every $z \in B_{ix}$ is μu for some $u \in B_x$, and $| B_x | = | B_{ix} |$. Since μ is a function it follows that $iy \in B_{ix}$. (3) is proved.

Proof of Theorem 1 from Theorem 2: Let R_2 be a fixed (i.e., nonrandomizing) optimal routing rule for ν_2, and define a fixed rule for ν_1 by the condition

$$r_{xy}^{(1)} = \begin{cases} 1, \text{ if } r_{(ix)(iy)}^{(2)} = 1 & \text{and} \qquad x, y \in M \\ 0, \text{ otherwise.} \end{cases}$$

Except for transient effects, use of R_1 will confine the trajectory of ν_1 to M, and gives the same blocking probability to ν_1 as use of R_2 yields for ν_2.

Acknowledgement

The author is indebted to S. Darlington and J. Riordan for valuable suggestions.

References

1. V. E. Beneš, Programming and Control Problems Arising from Optimal Routing in Telephone Networks, abstracted in *SIAM J. Control* 4 (1966), 6–18.

2. V. E. Beneš, *Mathematical Theory of Connecting Networks and Telephone Traffic*, Academic Press, New York, 1965.

3. G. Birkhoff, "Lattice Theory," *Am. Math. Soc. Colloq. Publ.* **25**, rev. ed., 1948, p. 3.